A GENERALIZED CLOSED FORM FOR TRIANGULAR MATRIX POWERS

Walter Shur

20 Speyside Circle, Pittsboro, North Carolina, 27312, USA

ABSTRACT.

[1] shows, for a $k\times k$ triangular matrix $M = [m_{i,j}]$ with unique diagonal elements, how to obtain numbers $p_{i,j,s}$ such that the $(i,j)^{th}$ element of M^n is given by ${}_nm_{i,j} = \sum_{s=i}^{j} p_{i,j,s} m_{s,s}^{n-1}$. This paper generalizes that formula: it removes the restriction of unique diagonal elements and shows how to obtain numbers $c_{i,j,r,s}$ such that ${}_nm_{i,j} = \sum_{r=1}^{num(i,j)} \sum_{s=1}^{mpy(r)} c_{i,j,r,s} \binom{n-1}{s-1} m_{i,j,r}^{n-s}$, where $num(i,j)$ is the number of unique diagonal elements between and including the i^{th} and j^{th} rows, $\{m_{i,j,r}\}_1^{num(i,j)}$ is the set of those unique elements, and $mpy(r)$ is the multiplicity of $m_{i,j,r}$ on that same range. Like the $p_{i,j,s}$ in [1], the $c_{i,j,r,s}$ are independent of the power to which the matrix is raised. The generalized formula works for any power of M, negative, zero or positive (positive only if the matrix is singular).

Key words: *Matrix, Triangular, Powers, Closed Form*

1. INTRODUCTION

[1] presents a method of obtaining a simple closed form for the powers of a triangular matrix with unique diagonal elements, as follows:

Definition 1. *Let $M = [m_{i,j}]$ be a $k \times k$ upper triangular matrix with unique diagonal elements. We define the power factors of M, $p_{i,j,s}$, recursively on the index j ,as follows:*

$$p_{i,j,s} = \frac{\sum_{t=s}^{j-1} p_{i,t,s} m_{t,j}}{m_{s,s} - m_{j,j}} \qquad i \le s < j \le k, \qquad \textit{(1.1)}$$

$$p_{i,j,s} = 0 \qquad s < i,\ s > j,$$

$$p_{i,j,j} = m_{i,j} - \sum_{t=i}^{j-1} p_{i,j,t} \qquad i < j \le k, \qquad \textit{(1.2)}$$

$$p_{j,j,j} = m_{j,j}.$$

Theorem 1. *If $M = [m_{i,j}]$ is a non-singular upper triangular matrix with unique diagonal elements, and ${}_nm_{i,j}$ is the $(i,j)^{th}$ element of M^n, then*

$$_nm_{i,j} = \sum_{s=i}^{j} p_{i,j,s} m_{s,s}^{n-1},$$

for all integral values of n, negative, positive or zero. If M is singular, the equation holds if $n \ge 1$ and 0^0 is taken as 1).

2. ALTERNATE DEFINITION FOR $p_{i,j,s}$.

Let $M = [m_{i,j}]$ be an upper triangular matrix with unique diagonal elements. The product

$$m_{i,a} m_{a,b} m_{b,c} \cdots m_{l,j},$$

where $i \le a < b < c \cdots < l < j$, and s is i, j, or one of a,b,c,$\cdots$ l, is called a chain from i to j passing through s. The length of the chain is the number of elements in the product. The expression

$$\frac{m_{i,a} m_{a,b} m_{b,c} \cdots m_{l,j}}{(m_{s,s} - m_{a,a})(m_{s,s} - m_{b,b})(m_{s,s} - m_{c,c}) \cdots (m_{s,s} - m_{l,l})(m_{s,s} - m_{j,j})}$$

where $(m_{s,s} - m_{s,s})$ is taken as 1, is called an adjusted chain from i to j passing through s.

Definition 2. *If $i \le s \le j$, $p_{i,j,s}$ is the sum of all adjusted chains from i to j passing through s. If $s < i$ or $s > j$, $p_{i,j,s} = 0$.*

Following are a few illustrative examples which help clarify the definition:

$$p_{1,1,1} = m_{1,1},$$

$$p_{1,3,1} = \frac{m_{1,1}m_{1,3}}{(m_{1,1}-m_{3,3})} + \frac{m_{1,1}m_{1,2}m_{2,3}}{(m_{1,1}-m_{2,2})(m_{1,1}-m_{3,3}),}$$

$$p_{1,3,2} = \frac{m_{1,2}m_{2,3}}{m_{2,2}-m_{3,3}} + \frac{m_{1,1}m_{1,2}m_{2,3}}{(m_{2,2}-m_{1,1})(m_{2,2}-m_{3,3}),}$$

$$p_{1,3,3} = m_{1,3} + \frac{m_{1,1}m_{1,3}}{m_{3,3}-m_{1,1}} + \frac{m_{1,2}m_{2,3}}{m_{3,3}-m_{2,2}} + \frac{m_{1,1}m_{1,2}m_{2,3}}{(m_{3,3}-m_{1,1})(m_{3,3}-m_{2,2}).}$$

Theorem 2. *Definition 2 is equivalent to Definition 1.*

Proof. From Definition 2, each term of the summand in (1.1) is of the form

$$\frac{m_{i,a}m_{a,b}m_{b,c}\cdots m_{l,t}m_{t,j}}{(m_{s,s}-m_{a,a})(m_{s,s}-m_{b,b})(m_{s,s}-m_{c,c})\cdots(m_{s,s}-m_{l,l})(m_{s,s}-m_{t,t})},$$

where $(m_{s,s} - m_{s,s})$ is taken as 1. The sum of all such terms, from $t = s$ to $t = j-1$, clearly includes all of the chains from i to j passing through s. That sum would be the sum of all the adjusted chains from i to j passing through s, except that the difference $(m_{s,s} - m_{j,j})$ is missing from each denominator. Hence the division by that difference in (1.1), and thus (1.1) is satisfied.

Next, we need to show that (1.2) is satisfied. The sum $\displaystyle\sum_{t=i}^{j} p_{i,j,t}$ consists of all adjusted chains from i to j of the form

$$\frac{m_{i,a_1}m_{a_1,a_2}m_{a_2,a_3}\cdots m_{a_r,j}}{(m_{s,s}-m_{a_1,a_1})(m_{s,s}-m_{a_2,a_2})(m_{s,s}-m_{a_3,a_3})\cdots(m_{s,s}-m_{a_r,a_r})(m_{s,s}-m_{j,j})},$$

with values $r = 1, 2, 3, \cdots, j-i$ and $s = a_1, a_2, a_3, \cdots, j$. The only term with $r = 1$ comes from $p_{i,j,j}$ and is equal to $m_{i,j}$ (see the illustrative

example for $p_{1,3,3}$ in Definition 2). The sum of all the adjusted chains of length 2 with the same numerator is

$$m_{i,a_1} m_{a_1,a_2} \left(\frac{1}{m_{a_1,a_1} - m_{a_2,a_2}} + \frac{1}{m_{a_2,a_2} - m_{a_1,a_1}} \right) = 0.$$

The sum of all adjusted chains of length 3 with the same numerator is

$$m_{i,a_1} m_{a_1,a_2} m_{a_2,a_3} \left(\frac{1}{(m_{a_1,a_1} - m_{a_2,a_2})(m_{a_1,a_1} - m_{a_3,a_3})} + \frac{1}{(m_{a_2,a_2} - m_{a_1,a_1})(m_{a_2,a_2} - m_{a_3,a_3})} + \frac{1}{(m_{a_3,a_3} - m_{a_1,a_1})(m_{a_3,a_3} - m_{a_2,a_2})} \right) = 0.$$

In general, the multiplier of $m_{i,a_1} m_{a_1,a_2} m_{a_2,a_3} \cdots m_{a_r,j}$ is seen [2] to be the $(r-1)^{st}$ divided difference of the polynomial $f(x) = 1$, and hence is 0 if $r \geq 2$.

Therefore, $\sum_{t=i}^{j} p_{i,j,t} = m_{i,j}$, and (1.2) is satisfied. And since $p_{j,j,j} = m_{i,j}$, Definition 2 is equivalent to Definition 1.

□

3. NON-UNIQUE DIAGONAL ELEMENTS

Theorem 3. *Let $M = [m_{i,j}]$ be an upper $k \times k$ triangular matrix with non-unique diagonal elements. Let $num(i,j)$ be the number of unique diagonal elements between and including the i^{th} and j^{th} rows, $\{m_{i,j,r}\}_1^{num(i,j)}$ be the set of those unique elements, and $mpy(r)$ be the multiplicity of $m_{i,j,r}$ on that same range ($\sum_{t=1}^{num(x,y)} mpy(r) = k$).*

Then

$${}_n m_{i,j} = \sum_{r=1}^{num(i,j)} \sum_{s=1}^{mpy(r)} c_{i,j,r,s} \binom{n-1}{s-1} m_{i,j,r}^{n-s},$$

where the $c_{i,j,r,s}$ are independent of the power to which the matrix is raised. This generalized formula works for any power of M, negative, zero or positive (positive only if the matrix is singular).

Proof. The proof is much easier to present, and much easier to follow, for a specific case. The generalization in Theorem 3 will be evident, including the method of determining the numerical values of $c_{i,j,r,s}$.

Let

$$M = \begin{pmatrix} 3 & 2 & 3 & 5 & 4 & 2 \\ 0 & 5 & 2 & 4 & 3 & 1 \\ 0 & 0 & 3 & 2 & 6 & 4 \\ 0 & 0 & 0 & 5 & 5 & 1 \\ 0 & 0 & 0 & 0 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

In terms of Theorem 3,

$$\begin{aligned} m_{1,6,1} &= 3 \quad mpy(1) = 3, \\ m_{1,6,2} &= 5 \quad mpy(2) = 2, \\ m_{1,6,3} &= 7 \quad mpy(3) = 1. \end{aligned}$$

and $num(1,6) = 3$. In order to make use of Theorem 1, we first alter the matrix M so that it will have unique diagonal elements, as follows:

$$M = \begin{pmatrix} 3 & 2 & 3 & 5 & 4 & 2 \\ 0 & 5 & 2 & 4 & 3 & 1 \\ 0 & 0 & 3-e & 2 & 6 & 4 \\ 0 & 0 & 0 & 5-f & 5 & 1 \\ 0 & 0 & 0 & 0 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 3-g \end{pmatrix}$$

At an appropriate stage, we will let e, f, and g become zero to obtain results for the original matrix M.

We have from Theorem 1, for example, the element of M^n at row 1 and column 6 is given by

$$
\begin{aligned}
{}_n m_{1,6} = p_{1,6,1}3^{n-1} + p_{1,6,2}5^{n-1} + p_{1,6,3}(3-e)^{n-1} + \\
p_{1,6,4}(5-f)^{n-1} + p_{1,6,5}7^{n-1} + p_{1,6,6}(3-g)^{n-1}.
\end{aligned}
\tag{3.1}
$$

Expanding the binomials, and collecting only those terms involving powers of 5, we obtain

$$
(p_{1,6,2} + p_{1,6,4})5^{n-1} + p_{1,6,4}\sum_{i=1}^{n-1}\binom{n-1}{i}(-f)^i 5^{n-1-i}. \tag{3.2}
$$

We know from Definition 2 that $p_{1,6,4}$ is the sum of all adjusted chains from 1 to 6 passing through 4. Combine all of those adjusted chains into a single fraction with the denominator equal to the product of the differences between $m_{4,4}$ and each of the other diagonal elements, namely $(2-f)(-f)(2+e-f)(-2-f)(2+g-f)$. We see from this denominator that $p_{1,6,4}(-f)^i = 0$ if $i > 1$ and $f = 0$. Therefore, the only terms of (3.2) involving powers of 5 are

$$
(p_{1,6,2} + p_{1,6,4})5^{n-1} + p_{1,6,4}(-f)\binom{n-1}{1}5^{n-2}.
$$

We determine the coefficients of the powers of 5, $c_{1,6,2,1}$ and $c_{1,6,2,2}$, using Mathematica as follows:

$c_{1,6,2,1}$ Input: $(p[1,6,2]+p[1,6,4]//\text{Together})/.\{e->0, f->0, g->0\}$

Output: $-\frac{59}{2}$

$c_{1,6,2,2}$ Input:$((-f)p[1,6,4]//\text{Together})/.\{e->0, f->0, g->0\}$

Output: -60

Similarly, collecting only those terms involving powers of 3, we obtain

$$(p_{1,6,1}+p_{1,6,3}+p_{1,6,6})3^{n-1}+\sum_{i=1}^{n-1}(p_{1,6,3}(-e)^i+p_{1,6,6}(-g)^i)\binom{n-1}{i}3^{n-1-i}. \tag{3.3}$$

Combine all of the adjusted chains in $(p_{1,6,3}(-e)^i + p_{1,6,6}(-g))^i$ into a single fraction, with the denominator equal to a x b, where a is the product of the differences between $m_{3,3}$ and each of the other diagonal elements, and b is the product of the differences between $m_{6,6}$ and each of the other diagonal elements. The only elements in that denominator which become zero when e, f, and g become zero are $(-e)(g-e)(e-g)(-g) = eg(g-e)^2$. If $(p_{1,6,3}(-e)^i + p_{1,6,6}(-g)^i)$ were equal to a constant $A \neq 0$, the numerator of that single fraction would have to contain a term $eg(g-e)^2A$. If $i > 2$, that is not possible since each term of the numerator must contain a power of e greater than 2 or a power of g greater than 2, but $eg(g-e)^2A$ contains the term $(-2)e^2g^2A$. Hence $(p_{1,6,3}(-e)^i + p_{1,6,6})(-g)^i = 0$ if $i > 2$ and e, f, and g are zero. Therefore the only terms of (3.3) involving powers of 3 are

$$(p_{1,6,1}+p_{1,6,3}+p_{1,6,6})3^{n-1} + ((-e)p_{1,6,3} + (-g)p_{1,6,6})\binom{n-1}{1}3^{n-2} + \\ ((-e)^2p_{1,6,3} + (-g)^2p_{1,6,6})\binom{n-1}{2}3^{n-3}.$$

Again, we determine the coefficients of powers of 3, $c_{1,6,1,1}, c_{1,6,1,2}, c_{1,6,1,3}$, using Mathematica as follows:

$c_{1,6,1,1}$ Input: $(p[1,6,1] + p[1,6,3] + p[1,6,6]//\text{Together})/.\{e->0, f->0, g->0\}$

Output: $-\frac{203}{32}$

$c_{1,6,1,2}$ Input: $((-e)p[1,6,3] + (-g)p[1,6,6]//\text{Together})/.\{e->0, f->0, g->0\}$

Output: $\frac{5}{8}$

$c_{1,6,1,3}$ Input: $((-e)^2 p[1,6,3] + (-g)^2 p[1,6,6]//\text{Together})/.\{e->0, f->0, g->0\}$

Output: $\frac{15}{2}$

The only term in (3.1) involving powers of 7 is $p_{1,6,5}7^{n-1}$. We determine the coefficient of 7^{n-1}, $c_{1,6,3,1}$, using Mathematica, as follows:

$c_{1,6,3,1}$ Input: (p[1,6,5]//Together)/.$\{e->0, f->0, g->0\}$

Output: $\frac{1211}{32}$

Combining the above results we obtain the closed form expression for ${}_nm_{1,6}$,

$$
{}_nm_{1,6} = -\frac{203}{32}3^{n-1} + \frac{5}{8}\binom{n-1}{1}3^{n-2} + \frac{15}{2}\binom{n-1}{2}3^{n-3} - \frac{59}{2}5^{n-1} - 60\binom{n-1}{1}5^{n-2} + \frac{1211}{32}7^{n-1}.
$$

In the terms of Theorem 3, this is

$$
{}_nm_{1,6} = c_{1,6,1,1}m_{1,6,1}^{n-1} + c_{1,6,1,2}\binom{n-1}{1}m_{1,6,1}^{n-2} + c_{1,6,1,3}\binom{n-1}{2}m_{1,6,1}^{n-3} + c_{1,6,2,1}m_{1,6,2}^{n-1} + c_{1,6,2,2}\binom{n-1}{1}m_{1,6,2}^{n-2} + c_{1,6,3,1}m_{1,6,3}^{n-1}.
$$

It is easy to see from the pattern in (3.1) that the structure of Theorem 3 is correct in the general case. To prove Theorem 3 we need only establish the cutoff points for the expansion of the binomials in (3.1). We did that explicitly for elements of multiplicity 2 and 3. We do so now for elements of higher multiplicities.

The argument immediately following (3.3) proved that for a diagonal element of multiplicity 3, the expression $(p_{1,6,3}(-e)^i + p_{1,6,6}(-g)^i)$ was

zero if $i > 2$ and e and g were zero. It did that by showing that the numerator would have to contain a term $eg(g-e)^2A(A \neq 0)$, which was then shown to be not possible.

Suppose now that the element $m_{5,5}$ in the unaltered matrix M were 3 instead of 7, and was replaced by $3-h$ in the altered matrix. That same argument, but now for a diagonal element of multiplicity 4, would change as follows:

(1) $(p_{1,6,3}(-e)^i + p_{1,6,6}(-g)^i)$ would be replaced by $(p_{1,6,3}(-e)^i + p_{1,6,5}(-h)^i + p_{1,6,6}(-g)^i)$,

(2) $eg(g-e)^2$ would be replaced by $egh(e-g)^2(e-h)^2(g-h)^2$,

(3) If $(p_{1,6,3}(-e)^i + p_{1,6,5}(-h)^i + p_{1,6,6}(-g)^i)$ were equal to a constant $A \neq 0$, the numerator would have to contain a term $egh(e-g)^2(e-h)^2(g-h)^2A$. But if $i > 3$, that would not be possible because each term of the numerator would have to contain a power greater that 3 for at least one of e, g, or h, but $egh(e-g)^2(e-h)^2(g-h)^2A$ contains the term $-8e^3g^3h^3A$.

The generalization to higher multiplicites is straightforward.

□

4. ILLUSTRATIONS

(1) We use Theorem 3 to obtain a closed form for ${}_nm_{2,4}$ for the matrix M in Section 3, noting that $m_{2,4,1} = 5$, $mpy(1) = 2$, $m_{2,4,2} = 3$, $mpy(2) = 1$, and $num(2,4) = 2$. The coefficients $c_{2,4,r,s}$, are determined using Mathematica as follows:

$c_{2,4,1,1}$ Input: $(p[2,4,2] + p[2,4,4]//\text{Together})/.\{e- > 0, f- > 0\}$

Output: 1

$c_{2,4,1,2}$ Input:$((-f)p[2,4,4]//\text{Together})/.\{e->0, f->0\}$

Output: 30

$c_{2,4,2,1}$ Input:$(p[2,4,3]//\text{Together})/.\{e->0, f->0\}$

Output: 3

Theorem 3 gives us

$$ {}_n m_{2,4} = 1 \cdot 5^{n-1} + 30 \binom{n-1}{1} 5^{n-2} + 3 \cdot 3^{n-1}. $$

(2) Let M and the altered M be the matrices

$$ \begin{pmatrix} 5 & 2 & 1 & 3 \\ 0 & 5 & 4 & 2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix} \text{and} \begin{pmatrix} 5 & 2 & 1 & 3 \\ 0 & 5-e & 4 & 2 \\ 0 & 0 & 5-f & 1 \\ 0 & 0 & 0 & 5-g \end{pmatrix} $$

We use Theorem 3 to obtain a closed form for ${}_n m_{1,4}$, noting that $m_{1,4,1} = 5$, $mpy(1) = 4$, and $num(1,4) = 1$. The coefficients $c_{1,4,r,s}$ are determined using Mathematica as follows:

$c_{1,4,1,1}$ Input: $(p[1,4,1] + p[1,4,2] + p[1,4,3] + p[1,4,4]//\text{Together})/.\{e->0, f->0\}, g->0\}$

Output: 3

$c_{1,4,1,2}$ Input:$((-e)p[1,4,2] + (-f)p[1,4,3] + p(-g)p[1,4,4]//\text{Together})/.\{e->0, f->0, g->0\}$

Output: 20

$c_{1,4,1,3}$ Input:$((-e)^2 p[1,4,2] + (-f)^2 p[1,4,3] + p(-g)^2 p[1,4,4]//\text{Together})/.\{e->0, f->0\, g->0\}$

Output: 33

$c_{1,4,1,4}$ Input:$((-e)^3 p[1,4,2] + (-f)^3 p[1,4,3] + p(-g)^3 p[1,4,4]//\text{Together})/.\{e->0, f->0\, g->0\}$

Output: 40

Theorem 3 gives us

$$ {}_n m_{1,4} = 3 \cdot 5^{n-1} + 20\binom{n-1}{1}5^{n-2} + 33\binom{n-1}{2}5^{n-3} + 40\binom{n-1}{3}5^{n-4}. $$

Note: In the above illustrations, three different variables (e,f and g) were chosen for clarity of presentation. In actual practice, replacing (e,f and g) with variables such as e, e^2, and e^3 would be more efficient since Mathematica would need to record only one variable instead of three, and only one variable need be set to zero.

References

[1] W. Shur. A Simple Closed Form For Triangular Matrix Powers. *Electronic Journal of Linear Algebra,* 22:1000-1003, 2011

[2] V. Jankovic. Divided Differences. *The Teaching Of Mathematics,* III, 2:115-119, 2000.

[3] C.P. Huang, An Efficient Algorithm For Computing Powers Of Triangular Matrices, *ACM '78 Proceedings of the 1978 Annual Conference,* 2:954-957, 1978

(Walter Shur, 919-542-7179, wrshur@gmail.com, 1/28/13)